AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 105:109–119 (1998)

mtDNA and Language Support a Common Origin of Micronesians and Polynesians in Island Southeast Asia

J. KOJI LUM[1]* and REBECCA L. CANN[2]
[1]*Department of Human Genetics, Eccles Institute of Human Genetics, University of Utah Health Sciences Center, Salt Lake City, Utah 84112-5330*
[2]*Department of Genetics, John A. Burns School of Medicine, University of Hawai'i at Manoa, Honolulu, Hawai'i 96822*

KEY WORDS region V; Remote Oceania; Melanesia; Austronesian; geographic isolation

ABSTRACT The origins and relationships among Micronesians, Polynesians, and Melanesians were investigated. Five different mtDNA region V length polymorphisms from 873 individuals representing 24 Oceanic and Asian populations were analyzed. The frequency cline of a common deletion and the distributions of a rare expanded length polymorphism support the origin of both Micronesians and Polynesians in Island Southeast Asia. Genetic, linguistic, and geographic distances were compared to assess the relative importance of isolation and gene flow during the prehistory of 19 Austronesian-speaking populations subdivided into five potential spheres of interaction. We observed significant correlations ($P < 0.05$) between genetic and linguistic distances in four of five comparisons. These data indicate extensive gene flow throughout much of Micronesia, but substantial isolation in other Pacific regions. Although recent advancements in our understanding of intentional voyaging within Remote Oceania have challenged the existence of the "myth of the primitive isolate," we caution against the adoption of panmictic alternatives. Am J Phys Anthropol 105:109–119, 1998.

Ever since European explorers ventured into the Pacific and found virtually every island inhabited, scholars have debated the origins and relationships among Micronesians, Polynesians, and Melanesians. Archaeological and linguistic evidence supports two distinct stages of Pacific Island settlement. The first phase resulted in the colonization of the Melanesian regions of New Guinea at least 40,000 years before present (ybp) (Groube et al., 1986), the Bismarck Archipelago 33,000 ybp (Allen et al., 1988), and the Northern Solomons 29,000 ybp (Wickler and Spriggs, 1988). The distribution of these Pleistocene sites coincides with the distribution of Papuan-speaking people in the Pacific.

The colonization of Remote Oceania (Pawley and Green, 1973; Green, 1991) is dated to the last 4,000 ybp. This population expansion is associated exclusively with Austronesian languages and includes the remaining areas of Melanesia and all of Polynesia and Micronesia. The most diverse Austronesian languages in Oceania are found in Western Micronesia. In particular, the languages of the Marianas Islands and Palau are not closely related to each other and are among the most distinct of all Austronesian lan-

Contract grant sponsor: Wenner-Gren Foundation for Anthropological Research; contract grant number 5640; Contract grant sponsor: NSF/Sloan; contract grant number BIR-9510748; Contract grant sponsor: NSF; contract grant number SBR-9514733; Contract grant sponsor: NSF; contract grant number SBR-9512178.

*Correspondence to: J. Koji Lum, Ph.D., University of Utah, Department of Human Genetics, Eccles Institute of Human Genetics, 15 North 2030 East Street, Room 2100, Salt Lake City, UT 84112-5330. Fax: (801) 581-7796. E-mail: koji@odin.genetics.utah.edu

Received 27 June 1997; accepted 20 October 1997.

guages (Bender, 1971). The linguistic classification of Yapese, the other language of Western Micronesia, has remained controversial (Ross, 1996). The languages of Eastern Micronesia, Remote Oceanic Melanesia, and Polynesia are classified within the Oceanic branch of the Austronesian-language family (Pawley and Ross, 1993).

In Melanesia and Polynesia the second wave of settlement is associated archaeologically with the Lapita cultural complex. This cultural complex is characterized by highly decorated red-slipped pottery and is found along the north coast of New Guinea, throughout Remote Oceanic Melanesia, and in Western Polynesia. Kirch and Hunt (1988) have argued that the spread of Lapita pottery through Melanesia and Western Polynesia was "sufficiently rapid so as to be undetectable within the statistical uncertainty range of radiocarbon dating." Diamond (1988) interprets this as evidence of a rapid expansion of people into Remote Oceania which he calls the "express train to Polynesia."

While there is general agreement about the expansion of the Lapita cultural complex east from the Bismarck Archipelago into Remote Oceania, there is considerable controversy surrounding its origin. Based on archaeological evidence and correlations with the distribution of Austronesian languages, some see it as evidence of a wave of emigration from Southeast Asia (Bellwood, 1989; Spriggs, 1989). In this scenario, the Austronesian-speaking passengers on the express train leave Island Southeast Asia and disperse throughout Oceania with little interaction with the descendants of the Pleistocene settlers of Western Melanesia. In contrast, Terrell (1989) sees Lapita as a set of cultural innovations by indigenous Melanesians resulting from long-term interactions with other Western Pacific populations.

These contrasting views result from a fundamental disagreement of the relative importance of isolation and interaction among Oceanic populations during prehistory. If isolation has characterized Oceanic prehistory, then each island can be thought of as a replicate "experiment" resulting from cultural fission and population dispersal. Under this model, correlated biological and cultural differences accumulate collinearly over time between geographically isolated populations. Thus we would expect biological relationships to mirror linguistic relationships which describe Micronesians, Polynesians, and Remote Oceanic Melanesians as more closely related to Austronesian-speaking Southeast Asians than to Papuan-speaking Melanesians.

If high rates of gene flow across linguistic boundaries were common, no significant correlation between genetic and linguistic patterns would be expected. Correlations between genetics and geography irrespective of linguistic relationships have been observed in comparisons of Austronesian- and Papuan-speaking Melanesians (Serjeantson et al., 1983; Welsch et al., 1992). Extrapolating from these studies, and bolstered by recent experimental (Lewis, 1972; Finney, 1994) and simulated (Irwin, 1992) voyaging, Terrell et al. (1997) have argued that extensive interactions have rendered linguistic patterns throughout the Pacific, including Remote Oceania, "uninformative."

To address questions of origins and interactions among Pacific Island populations we have generated length and sequence polymorphism data from region V of the mtDNA molecule in 873 individuals from nine Micronesian, four Polynesian, three Melanesian, and eight geographically or linguistically related populations. Length polymorphisms in region V of the mtDNA molecule were first detected in Asian and Polynesian individuals (Cann and Wilson, 1983; Wrischnik et al., 1987). Differences in the frequencies of region V length polymorphisms have been useful in elucidating population relationships among Asians (Horai and Matsunaga, 1986; Ballinger et al., 1992; Harihara et al., 1992; Passarino et al., 1993), Amerindians (Schurr et al., 1990; Shields et al., 1992; Lorenz and Smith, 1994), and Pacific Islanders (Hertzberg et al., 1989; Lum et al., 1994; Melton et al., 1995; Redd et al., 1995; Sykes et al., 1995). The presence of the 9-bp deletion associated with similar control region haplotypes at substantial frequencies in both Amerindian and Oceanic populations has also resurrected the question of

TABLE 1. Frequency of length polymorphisms and estimates of heterozygosity in Pacific Island and Asian populations

Population	Region[a]	Language[b]	n	XX[c]	X.I[d]	X.II[e]	XeX[f]	XXX[g]	h[h]
Rapa Nui	Pol	OcA	14		1.00				0.00
Kapingamarangi	Pol	OcA	35		1.00				0.00
Hawai'i	Pol	OcA	28	0.04	0.96				0.07
Marshalls	Mic	OcA	31	0.07	0.94				0.13
Outer Islands of Yap	Mic	OcA	137	0.09	0.91				0.16
Samoa	Pol	OcA	29	0.10	0.86		0.03		0.25
Nauru	Mic	OcA	28	0.11	0.89				0.20
Kosrae	Mic	OcA	30	0.20	0.80				0.33
Kiribati	Mic	OcA	27	0.26	0.74				0.40
Yap Proper	Mic	OcA	70	0.27	0.73				0.40
Fiji	Mel	OcA	14	0.29	0.64			0.07	0.54
Palau	Mic	A	134	0.42	0.56		0.01		0.51
Pohnpei	Mic	OcA	32	0.44	0.56				0.51
Philippines	ISEA	A	25	0.56	0.40		0.04		0.55
Vietnam	MA	O	25	0.64	0.32	0.04			0.51
Borneo	ISEA	A	21	0.57	0.38		0.05		0.55
Canton	MA	O	25	0.72	0.24	0.04			0.44
Java	ISEA	A	22	0.73	0.27				0.42
Marianas	Mic	A	53	0.87	0.13				0.23
Japan	MA	O	32	0.88	0.13				0.23
Vanuatu	Mel	OcA	25	0.88	0.12				0.22
Moken	IO	A	12	1.00					0.00
Papua New Guinea	Mel	P	12	1.00					0.00
Australian Aborigine	Aus	O	12	1.00					0.00
Total			873						

[a] Region abbreviations: Pol = Polynesia; Mic = Micronesia; Mel = Melanesia; ISEA = Island Southeast Asia; MA = Mainland Asia; Aus = Australia; IO = Indian Ocean.
[b] Language abbreviations: OcA = Oceanic Austronesian; A = Austronesian; P = Papuan; O = other.
[c] XX = two 9-bp repeats (ACCCCCTCT ACCCCCTCT).
[d] X.I = one 9-bp repeat (ACCCCCTCT).
[e] X.II = one 9-bp repeat (TCCCCCTCT).
[f] XeX = one expanded repeat and a 9-bp repeat (CCCCCCCCCCCCCC ACCCCCTCT).
[g] XXX = three 9-bp repeats (ACCCCCTCT ACCCCCTCT ACCCCCTCT).
[h] h = heterozygosity.

transpacific gene flow (Bonatto et al., 1996; Cann and Lum, 1996).

We identified five distinct length polymorphisms produced by deletions and duplications of perfect repeats, single base substitutions, and a slipped expansion of a homopolymeric tract. Analyses of these data suggest extensive gene flow within Micronesia, but substantial isolation among other groups of Austronesian-speaking populations of the Pacific.

MATERIALS AND METHODS

Subjects

This study was approved by the Internal Review Board of the University of Hawai'i at Manoa, the Scientific Review Board of the Federated States of Micronesia, and the traditional leaders of Palau (Rubekul Belau) and of Yap (Council of Pilung). Five plucked hairs were collected from 517 consenting participants. In many areas of Micronesia informed consent was obtained from participants with the assistance of the local hospital or clinic staff. An additional 356 DNA samples were obtained from other researchers. The study population consists of 24 groups from seven geographic regions (Table 1).

Methods

Genomic DNA was extracted from plucked hair samples using the method of Higuchi (1989) or the silica extraction protocol of Boom et al. (1990) abbreviated as follows. Five hairs were washed with ethanol and rinsed with sterile distilled water and the bulb ends were clipped into a sterile tube containing 200 µl of lysis buffer 6 (Boom et al., 1990) and incubated at 60°C for 2 hours. After centrifugation, the supernatant was transferred to a new tube, 20 µl of the silica solution was added, and the tube was vortexed. The sample was incubated for 10 minutes, then centrifuged to pellet the silica. The pellet was washed twice with 500 µl of 70% ethanol and once with 500 µl of acetone

and then loosely covered with foil and dried at 56°C. Genomic DNA was eluted from the dried pellet with 150 µl of sterile water. One microliter of the elutant was used to PCR amplify an approximately 121-bp fragment including the mtDNA region V using previously published primers (Wrischnik et al., 1987).

PCR products were separated by electrophoresis in either 6% (NuSieve, FMC) or 4% (2% NuSieve, FMC and 2% ultrapure agarose, Gibco) agarose gels and visualized by ethidium bromide staining. PCR products of non-standard lengths (i.e., other than two copies of the 9-bp repeat) were sequenced manually or on an ABI 377 automated sequencer. For these length polymorphisms, both the mtDNA heavy and light strands were sequenced.

Unbiased heterozygosity (h) for each population was estimated from frequencies of the five length polymorphisms using the equation (Nei, 1987):

$$h = (1 - \Sigma x^2)n/(n - 1)$$

Genetic distances between populations were calculated using the Cavalli-Sforza chord measure option of the GENDIST program of the PHYLIP version 3.5c analysis package (Felsenstein, 1993). This distance measure assumes no mutation and that differences of allele frequencies result from genetic drift, reasonable assumptions for the short time of settlement of Micronesia and Polynesia (<4,000 ybp).

Nineteen of the study populations speak Austronesian languages (Table 1). Based upon exclusively shared innovations (Bender, 1971; Pawley and Green, 1973; Jackson, 1983; Pawley and Ross, 1993) the relationships of the languages spoken by these populations appear in Figure 1. Yapese is considered an Oceanic language as argued by Ross (1996). Quantitative linguistic distances among these 19 populations were estimated by assigning relative branch lengths to the qualitative phylogeny. Geographic great circle distances (e.g. Spuhler, 1972) between populations were calculated from their latitudes and longitudes.

The genetic, linguistic, and geographic distance matrices were compared to each other using the Mantel permutation method

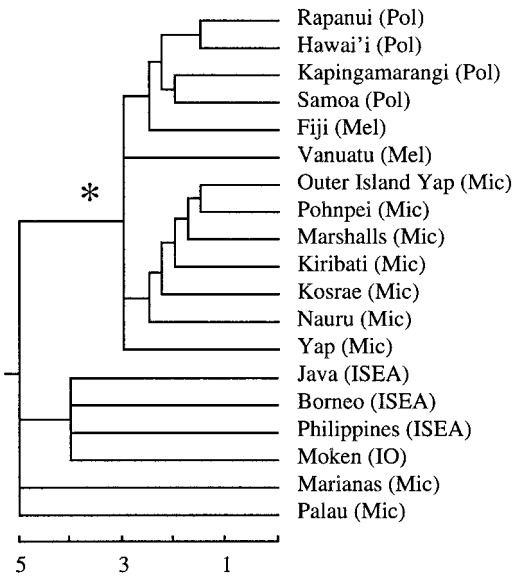

Fig. 1. Qualitative linguistic relationships among the 19 Austronesian-speaking populations studied. The asterisk leads to the 13 Polynesian, Melanesian, and Micronesian populations that speak Oceanic languages. The geographic region of each population is indicated in parentheses, the abbreviations are the same as in Table 1. Quantified linguistic distances between populations were estimated by assigning relative branch lengths to this tree. For example, the distance between Palau and any other population is estimated to be 10 units, and the distance between Yap and any other Oceanic-speaking population is estimated to be six units.

(Mantel, 1967) extended to examine partial correlations of multiple matrices (Smouse and Long, 1992) using Matrix Correlation Analysis version 1.0 (Long, 1996). To determine significance levels, the observed correlations were compared to a null distribution generated from 10,000 permutations. Multiple matrix comparisons were performed with five sets of Austronesian-speaking populations representing potential spheres of interaction: Oceanic-speaking Micronesians (seven populations), Polynesians and Melanesians (six populations), Oceanic speakers from Micronesia, Polynesia, and Melanesia (13 populations), all Micronesians (nine populations), and all speakers of Austronesian languages (19 populations).

RESULTS

Region V length polymorphisms

Five different length polymorphisms were observed among the 873 individuals. Fre-

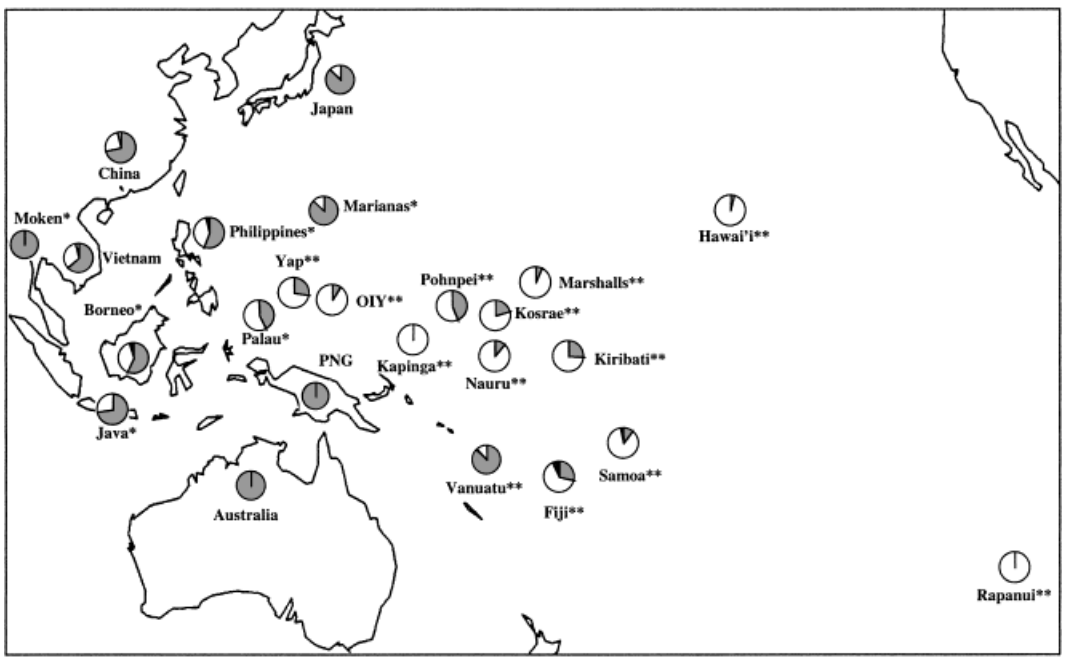

Fig. 2. Map displaying the frequencies of the region V length polymorphisms and geographic relationships of the nine Micronesian (Palau, Yap, OIY, Marianas, Pohnpei, Kosrae, Nauru, Marshalls, and Kiribati), four Polynesian (Kapingamarangi, Samoa, Hawai'i, and Rapanui), three Melanesian (PNG, Vanuatu, and Fiji) and eight other populations. OIY = Outer Island Yap; PNG = Papua New Guinea. The asterisk denotes Austronesian-speaking populations, while the double asterisks denote Oceanic Austronesian-speaking populations. In the pie charts: white = X.I; light grey = X.II; grey = XX; dark grey = XXX; and black = XeX (see Table 1 for numerical frequencies).

quencies of these length polymorphisms and the geographic relationships among the 24 populations are displayed in Figure 2. These frequencies along with population heterozygosities are also presented in Table 1.

The ancestral mtDNA region V consists of two copies of a 9-bp repeat (Horai et al., 1993), which is considered the "standard length" (XX). Possible mutational mechanisms producing the four non-standard length polymorphisms are depicted in Figure 3. The most common short-length polymorphism (X.I) results from the deletion of one 9-bp repeat as described by Wrischnik et al. (1987). X.I is found in all the populations except the Moken, Australian Aborigines, and Papuan-speaking Melanesians. Observed frequencies of X.I are low to moderate in Mainland Asia (0.12 to 0.32), moderate in Island Southeast Asia (0.27 to 0.40), high to near fixation in Micronesia (0.56 to 0.94) and near fixation in Polynesia (0.86 to 1.0). The exception to this pattern is the Marianas Islands in Northwestern Micronesia, which has a frequency of 0.13. There is a west to east cline in X.I frequencies in the Melanesian populations sampled. X.I is absent in Papua New Guinea, low in Vanuatu (0.13), and high in Fiji (0.64). All of the individuals with X.I share a phylogenetically distinct group of control region sequences (Cann and Lum, 1996) which correspond to Lineage Group I (Lum et al., 1995).

A second short length polymorphism (X.II) was detected in two individuals. X.II is most likely the result of two mutations (Fig. 3). The first mutation, a transition substitution 5′ to the first repeat, creates a novel, third reading frame of perfect 9-bp repeats. A deletion of a repeat in any of the three reading frames results in X.II, which is the same length as, but distinct in sequence from, the common short length polymorphism (X.I). Alternatively X.II could result from X.I via a single base substitution, but control region sequences from X.II do not

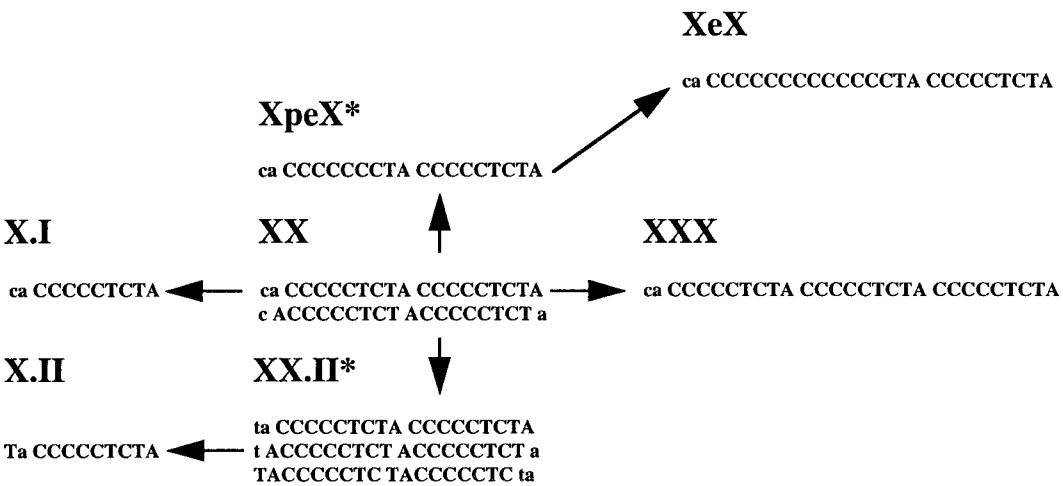

Fig. 3. Abbreviations and sequences of the five length polymorphisms related by inferred mutations. Types of mutations are represented by displacement in three directions. Horizontal arrows indicate deletions and duplications of perfect repeats, vertical arrows indicate single base substitutions, and the diagonal arrow indicates a replication slippage expansion of a homopolymeric tract. This diagram includes two putative "standard length" polymorphisms denoted by asterisks. The pre-expansion polymorphism (XpeX) contains a degenerate repeat including a homopolymeric tract of seven cytosines. XX contains nine base repeats in two reading frames while XX.II contains nine base repeats in three reading frames.

appear to be closely related to those of X.I (data not shown). This short length polymorphism is found only in Mainland Asia: China (0.04) and Vietnam (0.04). One of the two clusters of mtDNA haplotypes characterized by a 9-bp deletion and restricted to Vietnam reported by Ballinger et al. (1992) may be X.II.

One of the long length polymorphisms (XeX) is also most likely the result of two mutations (Fig. 3). A transition substitution in one repeat creates a seven base homopolymeric tract of cytosines. Following this substitution, the homopolymeric tract expands, presumably through replication slippage and expansion as described by Wrischnik et al. (1987). Individuals with XeX share a closely related set of control region sequences (data not shown). This length polymorphism is observed in Borneo (0.05), the Philippines (0.04), Palau (0.01), and Samoa (0.03).

The second long length polymorphism (XXX) consists of a duplication of a 9-bp repeat (Fig. 3). This length polymorphism was observed only in Fiji (0.07). XXX has been reported previously in one Chukchi individual from Siberia (Shields et al., 1992). It is unclear whether the sharing of this length polymorphism is a retention from a population ancestral to all Pacific Rim populations or two independent mutations. The possibility that a rare length polymorphism from an ancestral population could be retained in both Fiji and Siberia, while the presumably more common X.I was lost in Siberia, seems unlikely. It will be interesting to compare control region sequences from these two individuals to evaluate if this shared length polymorphism results from common descent or from parallel mutations that should be detected in haplotype analysis.

Matrix correlations

Results of the matrix correlations among Austronesian-speaking populations are presented in Table 2. No significant correlations were observed for the Oceanic-speaking Micronesian data set, consistent with extensive gene flow irrespective of linguistic relationships within much of Micronesia. In contrast, all of the remaining comparisons reveal significant correlations ($P < 0.05$) between genetic and linguistic distances when geographic distances are controlled (Table 2). These data suggest coevolution of genetic and linguistic patterns and substan-

TABLE 2. Pairwise and partial correlations of genetic, linguistic, and geographic distances of Austronesian-speaking populations

	Oceanic Micronesians (7 populations)	Polynesians and Melanesians (6 populations)	All Oceanic speakers (13 populations)	All Micronesians (9 populations)	All Austronesians (19 populations)
Pairwise correlations					
mtDNA × language	−0.417	0.793**	0.266*	0.553	0.361**
mtDNA × geography	−0.234	−0.127	0.064	−0.009	0.397*
Language × geography	0.234	−0.360	0.125	0.207	0.296*
Partial correlations[a]					
mtDNA × language (G)	−0.384	0.807**	0.261*	0.567*	0.278*
mtDNA × geography (L)	−0.154	0.28	0.032	−0.151	0.326*
Language × geography (M)	0.154	−0.43	0.112	0.254	0.178

* $P < 0.05$; ** $P < 0.01$.
[a] Letters in parentheses indicates distance held constant: G = geography; L = language; M = mtDNA.

tial isolation throughout the rest of the Pacific.

The comparisons including all 19 Austronesian-speaking populations reveal significant correlations ($P < 0.05$) among genetic, linguistic, and geographic distances. These are the only significant correlations including geographic distances. However, the correlation between linguistic and geographic distances is not significant when genetic distances are held constant. This indicates an interesting network in which linguistic distances are correlated with genetic distances and genetic distances are correlated with geographic distances but linguistic distances are not correlated with geographic distances. A nearly identical pattern of significant correlations resulted from comparisons using linguistic distances based upon cognate sharing (data not shown).

Since there is no recombination in the mitochondrial genome, it is expected that region V polymorphisms should correlate highly with polymorphisms in other parts of the mtDNA molecule. Indeed, the genetic distances between populations calculated from mtDNA region V length polymorphisms used for our comparisons are significantly correlated ($r = 0.902$, $P < 0.0001$) with distances based on mtDNA control region sequences from the same individuals.

DISCUSSION

We have detected five mtDNA region V length polymorphisms in Pacific Basin populations. Seven populations, including five island populations, have three of these length polymorphisms (Table 1). This is in contrast to most African, European, and Central Asian populations that lack all but the "standard length" polymorphism. The high heterozygosity of region V length polymorphisms in these populations may result from an elevation in the frequency of rare length polymorphisms through founder effects and genetic drift in subdivided and isolated island populations.

Although X.I is absent in Papua New Guinea (Table 1) we observe a moderate frequency of X.I throughout Island Southeast Asia (0.27 to 0.40). We also find consistently high frequencies of X.I in Micronesia (Table 1, Fig. 2), reaching near fixation in the Marshall Islands (0.94). As reported previously, the frequency of X.I also increases to near fixation in Polynesian populations (Hertzberg et al., 1989; Lum et al., 1994; Redd et al., 1995; Sykes et al., 1995). Similarly, XeX is found in Island Southeast Asia, Micronesia, and Polynesia (Table 1). Thus, consistently high frequencies of X.I throughout Remote Oceania and the distribution of XeX supports a common origin of both Micronesians and Polynesians in Island Southeast Asia.

Table 1 shows a qualitative association between language and genetics; populations that speak Oceanic Austronesian languages tend to have high frequencies of X.I. The exception to this trend is Vanuatu. As noted earlier, there is a west to east frequency cline in X.I across Melanesia. Of the two Austronesian-speaking Melanesian populations, Vanuatu in the west is closer to Papuan-speaking populations and has a relatively low frequency of X.I. Fiji, in contrast, is at the eastern edge of Melanesia and has a high X.I frequency, comparable to other Oce-

anic Austronesian-speaking populations in Micronesia (Table 1). Fijians sampled in this study are from the Yasawa Island Group in the extreme west of Fiji. In this population we observe an X.I frequency of 0.64 while Hertzberg et al. (1989) reported a frequency of 0.82 from Fiji. The difference in X.I frequencies may reflect a west to east gradient within Fiji analogous to that seen across Melanesia. We suspect that all Oceanic-speaking populations initially had high frequencies of X.I and interpret the low X.I frequency of Vanuatu as evidence of gene flow from Papuan-speaking Melanesia.

Studies of both α-globin haplotypes (Roberts-Thomson et al., 1996) and HLA allele frequencies (Serjeantson, 1985) have grouped Fijians and other Austronesian-speaking Melanesians with Papuan-speaking Melanesians to the exclusion of other Austronesian speakers. This suggests extensive gene flow east across Melanesia. Taken with the results from our mtDNA analysis, this indicates that gene flow from Papuan-speaking populations into Austronesian-speaking populations within Melanesia has been sex biased; gene flow suggested by studies of autosomal loci extends farther east than that observed in mitochondrial studies.

Terrell et al. (1997) have taken evidence of extensive gene flow between Papuan- and Austronesian-speaking populations within Melanesia and projected this pattern of interaction on the rest of Oceania. Our evaluation of the importance of isolation and interaction has focussed on Austronesian-speaking populations dispersed throughout the Pacific. We interpret significant correlations between genetic and linguistic distances as evidence of isolation between populations.

In all but one of our data sets, we observe significant correlations ($P < 0.05$) between genetic and linguistic distances when geographic distances are controlled (Table 2). Thus, linguistic patterns among Austronesian-speaking populations in general are predictive of genetic patterns. Significant correlations between genetic and linguistic distances are also found in Central Europe (Barbujani et al., 1994) and worldwide (Chen et al., 1995).

Our comparisons of genetic, linguistic, and geographic distances from Oceanic-speaking Micronesian populations reveal no significant correlations (Table 2). This is consistent with extensive interactions throughout much of Micronesia, in agreement with ethnographic information. The atolls of the central Carolines have maintained traditional navigational techniques and long distance voyaging in order to ameliorate environmental vulnerability (Ridgell et al., 1994). Traditionally these atolls were part of the Yap Empire, and yearly trade and exchange through the *sawei* system (Lingenfelter, 1975; Alkire, 1978) have provided opportunities for migration. In addition, Marck (1986) has shown that linguistic boundaries within Micronesia are coincident with the distance travelled during an overnight voyage, suggesting ongoing interactions throughout prehistory.

Genetic and linguistic distances from Polynesian and Melanesian populations are significantly correlated ($P < 0.01$). The high correlation coefficient (0.8) indicates collinear accumulation of genetic and linguistic distances and substantial isolation among populations. Polynesia is characterized by vast ocean distances, and the populations included here are among the most dispersed. Even though Kapingamarangi, a Polynesian outlier, is located geographically within Micronesia, there are ethnographic data indicating very limited interactions with neighboring populations (Alkire, 1978). Likewise Rapanui appears to have been settled once and then remained isolated (Metraux, 1940). Consistent with this isolation, interviews of the Tahitian navigator Tupaia by Cook during his first voyage revealed extensive knowledge of Central Polynesia, but not of Hawai'i, Rapanui, or the outliers (Lewis, 1972).

A significant correlation ($P < 0.05$) between genetic and linguistic distances is also observed when all Oceanic-speaking populations from Micronesia, Polynesia, and Melanesia are compared (Table 2). This indicates significant isolation among these populations even though extensive interaction within Micronesia was indicated. Early maps drawn by Carolinian navigators show nearly all the islands of Micronesia, but do not

include any from Polynesia, consistent with voyaging restricted to the northern Pacific (Hezel, 1983).

When Palau and the Marianas, the two non-Oceanic-speaking populations of Micronesia, are added to the Micronesian comparisons, the genetic and linguistic distances are significantly correlated when geographic distances are held constant (Table 2). This suggests that these two populations have been largely isolated from other Micronesian populations. Although there are early accounts of precontact voyages between the atolls of the central Carolines and the Marianas (Hezel, 1983) and prehistorical voyages from Yap to Palau to quarry stone money (Lingenfelter, 1975; Alkire, 1978), extensive female gene flow evidently did not occur.

The final set of comparisons includes all Austronesian-speaking populations. Significant pairwise correlations are observed among genetic, linguistic, and geographic distances, indicating substantial isolation (Table 2). Interestingly the correlation between linguistic and geographic distances is not significant when genetic distance is controlled. If the Pacific was colonized by a closely related group of Austronesian-speaking people, we may expect initially to find a significant correlation between genetic and linguistic distances and, due to the vast distances in Remote Oceania, a non-linear relationship between both of these and geographic distances.

In our comparisons, the linguistic and geographic distances are not significantly correlated with each other as expected, although the genetic distances are significantly correlated with both the linguistic and geographic distances. We interpret these results as support of a rapid colonization of Remote Oceania by a closely related group of Austronesian-speaking populations from Island Southeast Asia. Following this initial settlement extensive gene flow occurred between these recent migrants and neighboring people from Island Southeast Asia and Papuan-speaking Melanesians without major impact on the established linguistic relationships. The observed correlation between genetic and geographic distances in the large-scale comparisons is a result of this putative regional gene flow.

Our analyses are based upon maternally inherited mtDNA length polymorphisms. As discussed above, gene flow within Oceania appears to be sex biased. Analyses in progress including mtDNA control region sequence and autosomal microsatellite loci data from the same individuals will provide additional insights on settlement patterns and gene flow throughout the Pacific.

ACKNOWLEDGMENTS

We thank all the people who made this study possible by donating DNA samples. We also express our appreciation to those who facilitated the collection of samples: P. Acevedo, D. Addison, M. Andrew, S. Auerbach, R.K. Blaisdell, C. Ching, J. Clegg, G. Dever, A. di Piazza, J. Filefney, S. George, Y. Gibbons, J. Gilmatam, J. Haleyalmang, G. Heathcote, N. Hinshiranan, L. Humphreys, T. Hunt, Y. Hsia, S. Izutsu, A. Kugfas, J. Martinson, M. Mesngon, J. Mooteb, D. Otobed, N. Patrick, A. Polloi, E. Petrick, R. Puri, F. Rehuher, J. Scott, S. Serjeantson, D. Sol, J. Winter, and J. Wozniak. We also thank R. Blust and K. Rehg for their insights on Austronesian linguistics. Finally, we thank L.B. Jorde and M. Bamshad for helpful comments. This work was supported by grants from the Wenner-Gren Foundation for Anthropological Research (Gr. 5640), NSF/Sloan (BIR-9510748), NSF (SBR-9514733), and NSF (SBR-9512178).

REFERENCES

Alkire WH (1978) Coral Islanders. In W Goldschmidt (ed.): Worlds of Man: Studies in Cultural Ecology. Arlington Heights: AHM Publishing.

Allen J, Gosden C, Jones R, and White JP (1988) Pleistocene dates for the human occupation of New Ireland, northern Melanesia. Nature *331:*707–709.

Ballinger SW, Schurr TG, Torroni AT, Gan YY, Hodge JA, Hassan K, Chen K-H, and Wallace DC (1992) Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient Mongoloid migrations. Genetics *130:*139–152.

Barbujani G, Whitehead GN, Bertorelle G, and Nasidze IS (1994) Testing hypotheses on processes of genetic and linguistic change in the Caucasus. Hum. Biol. *66:*843–864.

Bellwood P (1989) The colonization of the Pacific: Some current hypotheses. In AV Hill and SW Serjeantson (eds.): The Colonization of the Pacific: A Genetic Trail. Oxford: Oxford Scientific Publications, Clarendon Press, pp. 1–59.

Bender BW (1971) Micronesian languages. In TA Sebeok (ed.): Current Trends in Linguistics. The Hague: Mouton, pp. 426–465.

Bonatto SL, Redd AJ, Salzano FM, and Stoneking M (1996) Lack of ancient Polynesian-Amerindian contact. Am. J. Hum. Genet. *59:*253–256.

Boom R, Sol CJA, Salimans MMM, Jansen CL, Wertheim-Van Dillen PME, and Van Der Noordaa J (1990) Rapid and simple method for purification of nucleic acids. J. Clin. Microbiol. *28:*495–503.

Cann RL and Lum JK (1996) Mitochondrial myopia: Reply to Bonatto et al. Am. J. Hum. Genet. *59:*256–258.

Cann RL and Wilson AC (1983) Length mutations in human mitochondrial DNA. Genetics *104:*699–711.

Chen J, Sokal RR, and Ruhlen M (1995) Worlwide analysis of genetic and linguistic relationships of human populations. Hum. Biol. *67:*595–612.

Diamond JM (1988) Express train to Polynesia. Nature *336:*307–308.

Felsenstein J (1993) PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.

Finney B (1994) Voyage of Rediscovery: A Cultural Odyssey Through Polynesia. Berkeley: University of California Press.

Green RC (1991) Near and Remote Oceania: Disestablishing "Melanesia" in cultural history. In A Pawley (ed.): Man and a Half: Essays in Pacific Anthropology and Ethnobotony in Honour of Ralph Bulmer. Auckland: The Polynesian Society, pp. 491–502.

Groube L, Chappell J, Muke J, and Price D (1986) A 40,000 year-old human occupation site at Huon Peninsula, Papua New Guinea. Nature *324:*453–455.

Harihara S, Hirai M, Suutou Y, Shimizu K, and Omoto K (1992) Frequency of a 9-bp deletion in the mitochondrial DNA among Asian populations. Hum. Biol. *64:* 161–166.

Hertzberg M, Mickleson KNP, Serjeantson SW, Prior JF, and Trent RJ (1989) An Asian-specific 9-bp deletion of mitochondrial DNA is frequently found in Polynesians. Am. J. Hum. Genet. *44:*504–510.

Hezel FX (1983) The first taint of civilization: A history of the Caroline and Marshall Islands in pre-colonial days, 1521–1885. Honolulu: University of Hawai'i Press.

Higuchi R (1989) Simple and rapid preparation of samples for PCR. In HA Erlich (ed.): PCR Technology. New York: Stockton Press, pp 31–43.

Horai S and Matsunaga E (1986) Mitochondrial DNA polymorphism in Japanese. II. Analysis with restriction enzymes of four or five base pair recognition. Hum. Genet. *72:*105–117.

Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, and Tajima K (1993) Peopling of the Americas, founded by four major lineages of mitochondrial DNA. Mol. Biol. Evol. *10:*23–47.

Irwin G (1992) The Prehistoric Exploration and Colonization of the Pacific. Cambridge: Cambridge University Press.

Jackson FH (1983) The internal and external relationships of the Trukic languages of Micronesia. Ph.D. dissertation, Honolulu, University of Hawai'i, Manoa.

Kirch PV and Hunt TL (1988) Radiocarbon dates from the Mussau Islands and the Lapita colonization of the Southwest Pacific. R.A.C.A.A. *30:*161–169.

Lewis D (1972) We, the Navigators. Honolulu: University of Hawai'i Press.

Lingenfelter SG (1975) Yap: Political Leadership and Culture Change in an Island Society. Honolulu: University of Hawai'i Press.

Long JC (1996) Matrix Correlation Analysis, version 1.0. Distributed by the author. Section on Population Genetics and Linkage, Laboratory of Neurogenetics, NIAAA/NIH, Rockville, Maryland.

Lorenz J and Smith DG (1994) Distribution of the 9-bp mitochondrial DNA region V deletion among North American Indians. Hum. Biol. *66:*777–788.

Lum JK, Ricards O, Ching C, and Cann RL (1994) Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. Hum. Biol. *66:*567–590.

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res. *27:* 209–220.

Marck JC (1986) Micronesian dialects and the overnight voyage. J. Pol. Soc. *95:*253–258.

Melton T, Peterson R, Redd AJ, Saha N, Sofro ASM, Martinson J, and Stoneking M (1995) Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. Am. J. Hum. Genet. *57:*403–414.

Metraux A (1940) Ethnology of Easter Island. Bernice Pauahi Bishop Museum Bull. *160*.

Nei M (1987) Molecular Evolutionary Genetics. New York: Columbia University Press.

Passarino G, Semino O, Modiano G, and Santachiara-Benerecetti AS (1993) COII/tRNA^{Lys} intergenic 9-bp deletion and other mtDNA markers clearly reveal that the Tharus (Southern Nepal) have Oriental affinities. Am. J. Hum. Genet. *53:*609–618.

Pawley A and Green R (1973) Dating the dispersal of the Oceanic languages. Ocean. Ling. *12:*1–67.

Pawley A and Ross M (1993) Austronesian historical linguistic and cultural history. Annu. Rev. Anthropol. *22:*425–459.

Redd AJ, Takezaki N, Sherry ST, McGarvey ST, Sofro ASM, and Stoneking M (1995) Evolutionary history of the COII/tRNA^{Lys} intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. Mol. Biol. Evol. *12:*604–615.

Ridgell R, Ikea M, and Uruo I (1994) The persistence of Central Carolinian navigation. ISLA: A Journal of Micronesian Studies *2:*181–206.

Roberts-Thomson JM, Martinson JJ, Norwich JT, Harding RM, Clegg JB, and Boettcher B (1996) An ancient common origin of Aboriginal Australians and New Guinea Highlanders is supported by &gr;-globin haplotype analysis. Am. J. Hum. Genet. *58:*1017–1021.

Ross M (1996) Is Yapese Oceanic? In B Nothofer (ed.): Reconstruction, Classification, Description. Festschrift in Honor of Isidore Dyen. Hamburg: Verlag Meyer, pp. 121–166.

Schurr TG, Ballinger SW, Gan Y-Y, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, Weiss KM, and Wallace DC (1990) Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies, suggesting they derived from four primary maternal lineages. Am. J. Hum. Genet. *46:*613–623.

Serjeantson SW (1985) Migration and admixture in the Pacific: Insights provided by human leucocyte antigens. In R Kirk and E Szathmary (eds.): Out of Asia: Peopling the Americas and the Pacific. Canberra: The Journal of Pacific History, Inc., pp. 133–145.

Serjeantson SW, Kirk RL, and Booth PB (1983) Linguistic and genetic differentiation in New Guinea. J. Hum. Evol. *12:*77–92.

Shields GF, Hecker K, Voevoda MI, and Reed J (1992) Absence of the Asian-specific region V mitochondrial marker in native Beringians. Am. J. Hum. Genet. *50:*758–765.

Smouse PE and Long JC (1992) Matrix correlation analysis in anthropology and genetics. Yrbk. Phys. Anthropol. *35:*187–213.

Spriggs M (1989) The dating of the Island Southeast Asia Neolithic: An attempt at chronometric hygiene and linguistic correlation. Antiquity *63:*587–613.

Spuhler JN (1972) Genetic, linguistic, and geographical distances in native North America. In JS Weiner and J Huizinger (eds.): The Assessment of Population Affinity in Man. Oxford: Oxford University Press, pp. 72–95.

Sykes B, Leiboff A, Low-Beer J, Tetzner S, and Richards M (1995) The origins of the Polynesians: An interpretation from mitochondrial lineage analysis. Am. J. Hum. Genet. *57:*1463–1475.

Terrell J (1989) Commentary: What Lapita is and what Lapita isn't. Antiquity *63:*623–626.

Terrell JE, Hunt TL, and Gosden C (1997) The dimensions of social life in the Pacific. Curr. Anthropol. *38:*155–195.

Welsch RL, Terrell J, and Nadolski JA (1992) Language and culture on the north coast of New Guinea. Am. Anthropol. *94:*568–600.

Wickler S and Spriggs M (1988) Pleistocene human occupation of the Solomon Islands, Melanesia. Antiquity *62:*703–706.

Wrischnik LA, Higuchi RH, Stoneking M, Erlich HA, Arnheim N, and Wilson AC (1987) Length mutations in human mitochondrial DNA: Direct sequencing of enzymatically amplified DNA. Nucleic Acids Res. *15:* 529–542.